2. 正定値カーネルと再生核ヒルベルト空間

正定値カーネルによるデータ解析
ー カーネル法の基礎と展開 ー

福水健次
統計数理研究所／総合研究大学院大学

統計数理研究所　公開講座
2011年1月13,14日

概要

- 正定値カーネル

  定義と例

- 再生核ヒルベルト空間と正定値カーネル

  再生核ヒルベルト空間の定義

  正定値カーネルとの関係（Moore-Aronszajn theorem）

  特徴写像

- 正定値カーネルとRKHSの基本的性質

  正定値カーネルの性質

  再生核ヒルベルト空間の性質

  Bochnerの定理

概要

- 正定値カーネル
  - 定義と例
- 再生核ヒルベルト空間と正定値カーネル
  - 再生核ヒルベルト空間の定義
  - 正定値カーネルとの関係（Moore-Aronszajn theorem）
  - 特徴写像
- 正定値カーネルとRKHSの基本的性質
  - 正定値カーネルの性質
  - 再生核ヒルベルト空間の性質
  - Bochnerの定理

正定値カーネル

- (実)正定値カーネル

Ω: 集合.　$k : \Omega \times \Omega \to \mathbf{R}$

$k(x,y)$ がΩ上の(実)正定値カーネルであるとは，次の2つを満たすことをいう

1．(対称性)　$k(x,y) = k(y,x)$

2．(正定値性)　任意の自然数 n とΩ の点 $x_1, \ldots, x_n$ に対し，

$$\left(k(x_i, x_j)\right)_{i,j=1}^{n} = \begin{pmatrix} k(x_1,x_1) & \cdots & k(x_1,x_n) \\ \vdots & \ddots & \vdots \\ k(x_n,x_1) & \cdots & k(x_n,x_n) \end{pmatrix}$$

(Gram行列)

が(半)正定値．すなわち，任意の実数 $c_1,\ldots, c_n$ に対し，

$$\sum_{i,j=1}^{n} c_i c_j k(x_i, x_j) \geq 0$$

- 複素数値の正定値カーネル

$$k:\Omega\times\Omega\to\mathbf{C}$$

の場合にも正定値性は以下のように定義される.

任意の自然数 n と, 任意のΩ の点 $x_1, \ldots, x_n$ と, 任意の複素数 $c_1,\ldots, c_n$ に対し, $\sum_{i,j=1}^{n} c_i \bar{c}_j k(x_i, x_j) \geq 0$ （$\bar{c}_j$ は複素共役）

が成り立つとき, $k(x,y)$ を正定値カーネルという.

– 上の条件から Hermite性 $k(y,x)=\overline{k(x,y)}$ は自然に従う.
（Exercise (やや難)）

正定値カーネルの例

- Euclid内積

$\Omega = \mathbf{R}^m$

$$k(x, y) = x^T y$$

- 多項式カーネル

$\Omega = \mathbf{R}^m$

$$k_P(x, y) = (x^T y + c)^d$$

（ d: 自然数, $c \geq 0$ ）

- ガウスカーネル（RBFカーネル)

$\Omega = \mathbf{R}^m$

$$k_G(x, y) = \exp\left(-\frac{1}{\sigma^2} \|y - x\|^2 \right)$$

（ $\sigma > 0$ ）

- Laplacianカーネル

$\Omega = \mathbf{R}^m$

$$k_L(x, y) = \exp\left(-\beta \sum_{i=1}^m |x_i - y_i| \right)$$

（ $\beta > 0$ ）

  – 正定値であることはあとでチェックする

内積と正定値カーネル

Proposition 2.1

Ωを集合，V を内積 (,)を持つベクトル空間とし，写像

$$\Phi : \Omega \to V$$

が与えられているとする．

このとき

$$k(x, y) = \left(\Phi(x), \Phi(y)\right)$$

で定義されるカーネルは正定値である．

証明(実の場合)

$$\sum_{i,j=1}^{n} c_i c_j k(x_i, x_j) = \sum_{i,j=1}^{n} c_i c_j \left(\Phi(x_i), \Phi(x_j)\right)$$

$$= \left(\sum_{i=1}^{n} c_i \Phi(x_i), \sum_{j=1}^{n} c_j \Phi(x_j)\right) = \left\| \sum_{i=1}^{n} c_i \Phi(x_i) \right\|^2 \geq 0$$

– 特徴ベクトルの内積は，正定値カーネルを与える．

概要

- 正定値カーネル
  - 定義と例
- 再生核ヒルベルト空間と正定値カーネル
  - 再生核ヒルベルト空間の定義
  - 正定値カーネルとの関係（Moore-Aronszajn theorem）
  - 特徴写像
- 正定値カーネルとRKHSの基本的性質
  - 正定値カーネルの性質
  - 再生核ヒルベルト空間の性質
  - Bochnerの定理

ヒルベルト空間の復習

- (実)ヒルベルト空間
  ベクトル空間で，内積 $\langle \;,\; \rangle$ が与えられており，
  完備性(コーシー列が常に収束する)を満たす．

  – ヒルベルト空間は，無限次元でもよい

  – 例） $L^2(a,b)$: 区間 (a,b) 上の2乗可積分関数全体

  $$\text{内積} \quad \langle f, g \rangle_{L^2(a,b)} = \int_a^b f(x)g(x)dx$$

  – ユークリッド空間 $\mathbf{R}^m$ もヒルベルト空間のひとつ．

  – 複素ヒルベルト空間も定義される．

再生核ヒルベルト空間

Def. 集合 Ω 上の再生核ヒルベルト空間(Reproducing kernel Hilbert space, RKHS) H とは，Ω 上の関数からなるヒルベルト空間であって，任意の $x \in \Omega$ に対し $\phi_x \in H$ があって，

$$\langle f, \phi_x \rangle = f(x) \qquad (\forall f \in H) \qquad \text{（再生性）}$$

を満たすものをいう.

– 再生核:
$$k(y,x) := \phi_x(y)$$

このとき
$$\langle f, k(\cdot,x) \rangle = f(x)$$

– 再生核は存在すれば一意的. (証明略 [Exercise])

– 再生核ヒルベルト空間は「関数空間」. しかしL^2とは大きく異なる.

注）$k(\cdot, x)$： x を固定した第1変数を変数とする1変数関数

再生核ヒルベルト空間と正定値カーネル

Theorem 2.2
再生核ヒルベルト空間 H の再生核 k は正定値カーネルである.

証明 (実の場合)

対称性:

$$k(x, y) = \langle k(\cdot, y), k(\cdot, x) \rangle = \langle k(\cdot, x), k(\cdot, y) \rangle = k(y, x).$$

正定値性:

$$\Phi(x) := \phi_x = k(\cdot, x)$$

と定義すると, 再生性により

$$k(x, y) = \langle k(\cdot, x), k(\cdot, y) \rangle = \langle \Phi(x), \Phi(y) \rangle$$

Prop. 1 を使え.

再生核ヒルベルト空間と正定値カーネル

Theorem 2.3 (Moore-Aronszajn)

$k(x,y)$:　集合 Ω 上の正定値カーネル

$\Rightarrow$

Ω 上の関数からなるヒルベルト空間 H_k で次の3つを満たすものが一意に存在する.

(1)　$k(\cdot,x) \in H_k$　　$(x \in \Omega$ は任意に固定)

(2)　有限和　$f = \sum_{i=1}^{n} c_i k(\cdot, x_i)$　の形の元は H_k の中で稠密

(3)（再生性）

$$f(x) = \langle f, k(\cdot, x) \rangle \qquad \forall f \in H_k,\ x \in \Omega$$

証明略（福水 2010 参照）

RKHS，正定値カーネル，特徴写像

- 正定値カーネルとRKHSは1対1に対応する

$$k \quad \Longleftrightarrow \quad H_k$$

1対1

- 特徴写像
  - カーネルトリック

$$\langle \Phi(x), \Phi(y) \rangle = k(x, y)$$

が成り立つためには，正定値カーネル k を用意し，

$$\Phi : \Omega \to H_k, \quad \Phi(x) = k(\cdot, x)$$

を考えれば十分.

  - 以降，常に上の特徴写像を考える.

再生核ヒルベルト空間の例

- 線形カーネル
  - RKHS for $k(x,y) = x^T y$

$$H = \{ f : \mathbf{R}^m \to \mathbf{R} \mid \exists a \in \mathbf{R}^m, f(u) = k(u,a) = a^T u \}$$

Inner product

$$\left\langle a^T u, b^T u \right\rangle = a^T b \qquad \left(\left\langle f_a, k(\cdot, b) \right\rangle = f_a(b) \right)$$

  - H はEuclidean空間とヒルベルト空間として同型.

$$H \cong \mathbf{R}^m, \quad a^T u \leftrightarrow a$$

  - 特徴写像は上の同型対応

線形カーネルは，ユークリッド空間の普通の内積に他ならない

- 有限集合 $\{1, 2, \ldots, m\}$ 上のRKHS
  - 正定値カーネル $K(i, j)$ ＝ 半正定値行列 K
  - 固有値分解 $$K = U \Lambda U^T$$
  $$U = (u_1, \ldots, u_m), \quad \Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_m)$$
  - 仮定： K は（狭義）正定値. *i.e.*, $\lambda_i > 0$.
  - RKHS：
  $$H = \{ f : \{1, \ldots, m\} \to \mathbf{R} \} = \{ (f(1), \ldots, f(m)) \in \mathbf{R}^m \}$$
  内積
  $$f = \sum_{i=1}^m a_i u_i, \quad g = \sum_{i=1}^m b_i u_i$$
  $$\boxed{\langle f, g \rangle = f^T K^{-1} g = \sum_{i=1}^m \frac{a_i b_i}{\lambda_i}}$$

– 再生性のチェック

$K = (k_1, ..., k_m)$（列ベクトル） $k_i = K(\cdot, i)$

任意の $f \in H$ に対し,

$$\left\langle f, K(\cdot, i) \right\rangle = f^T K^{-1} k_i = f^T K^{-1} k_i = f^T e_i = f(i)$$

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$$

- Gaussian Kernel
  – RKHS は無限次元.
  – 後で詳述する.

概要

- 正定値カーネル
  定義と例

- 再生核ヒルベルト空間と正定値カーネル
  再生核ヒルベルト空間の定義
  正定値カーネルとの関係（Moore-Aronszajn theorem）
  特徴写像

- 正定値カーネルとRKHSの基本的性質
  正定値カーネルの性質
  再生核ヒルベルト空間の性質
  Bochnerの定理

正定値性を保つ演算

Proposition 2.4 $k_i(x, y)$: Ω 上の正定値カーネル($i = 1,2,\ldots$).
以下も正定値カーネル.

(1) Positive combination $a_1 k_1(x,y) + a_2 k_2(x,y)$ $(a_1, a_2 \geq 0)$

(2) Product $k_1(x,y) k_2(x,y)$

(3) Limit $k(x,y) = \lim_{n\to\infty} k_n(x,y)$

正定値カーネルのなす凸錐
ある集合 Ω 上の正定値カーネル全体は，積について閉じた
(各点位相での)閉凸錐

証明)

1)は自明. 3)は正定値性の条件が極限操作で保たれることからわかる.

2)を示すには, 2つの半正定値行列 A, Bの成分ごとの積(Hadamard積)がまた半正定値であることを示せばOK.

<u>Lemma 2.5</u>.

2つの半正定値行列 A, B のHadamard積は半正定値である.

証明)

$A = \sum_a \lambda_a u_a u_a^T$　:　固有値分解(λ_a 固有値, u_a 固有ベクトル)

$$\sum_{i,j=1}^n c_i c_j A_{ij} B_{ij} = \sum_{a=1}^n \sum_{i,j=1}^n c_i c_j \lambda_a u_a^i u_a^j B_{ij}$$

$$= \sum_{a=1}^n \lambda_a \left(\sum_{i,j=1}^n \boxed{c_i u_a^i} \boxed{c_j u_a^j} B_{ij} \right) \geq 0$$

(半正定値性より $\lambda_a \geqq 0$)

正定値カーネルの例：再論（証明）

- 多項式カーネル

$x^T y$ は正定値

$\Rightarrow$ $x^T y + c$ は正定値（ $c \geqq 0$ ）

$\Rightarrow$ $(x^T y + c)^d$ は正定値（d 個の積）

- ガウスカーネル

$x^T y$ は正定値

$\Rightarrow$ $e^{\frac{1}{\sigma^2} x^T y} = 1 + \frac{1}{\sigma^2} x^T y + \frac{1}{2!}\left(\frac{1}{\sigma^2} x^T y\right)^2 + \frac{1}{3!}\left(\frac{1}{\sigma^2} x^T y\right)^3 + \cdots$ は正定値

$\Rightarrow$ $e^{-\frac{1}{2\sigma^2} x^T x} e^{\frac{1}{\sigma^2} x^T y} e^{-\frac{1}{2\sigma^2} y^T y} = \exp\left(-\frac{1}{\sigma^2} \| x - y \|^2\right)$ は正定値

(Prop. 2.6)

正定値性を保つ変換

Proposition 2.6.

$k(x,y)$: Ω 上の正定値カーネル,　$f:\Omega\to\mathbf{R}$:任意の関数とすると,

$$\tilde{k}(x,y)=f(x)k(x,y)f(y)$$

は正定値.

特に,正定値カーネルが $k(x,x)>0$ ($x\in\Omega$ は任意)を満たすとき,

$$\tilde{k}(x,y)=\frac{k(x,y)}{\sqrt{k(x,x)k(y,y)}}=\frac{\langle\Phi(x),\Phi(y)\rangle}{\|\Phi(x)\|_{H_k}\|\Phi(y)\|_{H_k}}$$

は正定値(normalization).

証明）略 [Exercise]

例）

$$k(x,y)=(x^Ty+c)^d \quad (\ c>0\ ) \quad\Longrightarrow\quad \tilde{k}(x,y)=\frac{(x^Ty+c)^d}{(x^Tx+c)^{d/2}(y^Ty+c)^{d/2}}$$

和・積に対するRKHS

$(H_1, k_1), (H_2, k_2)$ ：RKHS とその正定値カーネル

- 和空間 $H_1 + H_2$
  - 正定値カーネルの和 $k = k_1 + k_2$ に対応するRKHSは，ベクトル空間として
    $$H_1 + H_2 = \{ f : \Omega \to \mathbf{R} \mid \exists f_1 \in H_1, \exists f_2 \in H_2 \text{ such that } f = f_1 + f_2 \}$$
    - 内積 $$\| f \|^2 = \min_{f = f_1 + f_2,\, f_1 \in H_1,\, f_2 \in H_2} \| f_1 \|_{H_1}^2 + \| f_2 \|_{H_2}^2$$

- 積空間 $H_1 \otimes H_2$
  - 正定値カーネルの積 $k = k_1 k_2$ に対応するRKHSは，ベクトル空間としてテンソル積.
    $f = \sum_{i=1}^n f_i g_i \quad \left(f_i \in H_1, g_i \in H_2 \right)$ の形の関数が稠密.
    - 内積
    $$\left\langle \sum_{i=1}^n f_i^{(1)} g_i^{(1)}, \sum_{j=1}^m f_j^{(2)} g_j^{(2)} \right\rangle = \sum_{i=1}^n \sum_{j=1}^m \left\langle f_i^{(1)}, f_j^{(2)} \right\rangle_{H_1} \left\langle g_i^{(1)}, g_j^{(2)} \right\rangle_{H_2}$$

連続なカーネルに対するRKHS

– 連続性／微分可能性

位相空間Ω上の正定値カーネル $k(x, y)$ が連続のとき,
定義されるRKHS H_k の関数はすべて連続関数.

$$\because)\quad |f(x)-f(y)|^2 = |\langle f, k(\cdot,x)-k(\cdot,y)\rangle|^2 \le \|f\|^2 \|k(\cdot,x)-k(\cdot,y)\|^2$$
$$= \|f\|^2 (k(x,x)-2k(x,y)+k(y,y)) \to 0 \quad (x \to y)$$

実は, $\mathbf{R}^m$ 上の正定値カーネル $k(x,y)$ が微分可能のとき,
H_k のすべての関数が微分可能.

平行移動不変なカーネル

Def.　$\mathbf{R}^m$ 上の正定値カーネル $k(x,y)$ が平行移動不変(shift-invariant)であるとは，ある1変数関数 $\phi(z)$ があって，$k(x,y) = \phi(x-y)$ と書けることをいう.

Ex)　Gauss カーネル，ラプラスカーネル

– Fourier kernel (complex-valued):

$$\exp\left(\sqrt{-1}\omega^T(x-y)\right) = \exp\left(\sqrt{-1}\omega^T x\right)\exp\left(-\sqrt{-1}\omega^T y\right) \quad \Rightarrow \text{ 正定値}$$

– $k(x,y) = \phi(x-y)$が正定値カーネルとなるとき，$\phi(z)$ を正定値関数という.

- ガウスカーネルの正定値性(別証明)
  - ガウス関数の逆Fourier変換表示

$$\exp\left(-\|x\|^2/2\right) = \frac{1}{(2\pi)^m}\int_{R^m}\exp\left(-\|\omega\|^2/2\right)\exp\left(\sqrt{-1}\omega^T x\right)d\omega$$

ガウス関数の Fourier変換は，またガウス関数

  - 正定値性の証明

$$\sum_t \underbrace{\exp\left(-\|\omega_t\|^2/2\right)}_{\text{正の数}}\underbrace{\exp\left(\sqrt{-1}\omega_t^T(x-y)\right)}_{\text{正定値カーネル}} \quad \to \quad \underbrace{\exp\left(-\|x-y\|^2/2\right)}_{\text{正定値}}$$

Bochnerの定理

Theorem 2.7（Bochner)

$\mathbf{R}^m$ 上の連続関数$\phi(z)$ が正定値であるための必要十分条件は，$\mathbf{R}^m$ 上のある非負有限Borel測度 Λ が存在して

$$\phi(z) = \int \exp\left(\sqrt{-1}\omega^T z\right) d\Lambda(\omega)$$

が成り立つことである.

積分は Lebesgue-Stieltjes 積分.

– 十分性： Λが連続な密度関数 $\rho(\omega)$ を持つ場合は

$$\sum_t \exp\left(\sqrt{-1}\omega_t^T z\right)\rho(\omega_t) \to \phi(z)$$

必要性は少し難しい*.

* さまざまな証明が知られている．参考文献は福水2010参照.

RKHSのFourier変換による表示

仮定： Bochnerの表示で，Λが連続な密度関数 $\rho(\omega)$ を持つとせよ.

$$k(x,y) = \int \exp\left(\sqrt{-1}\omega^T(x-y)\right)\rho(\omega)d\omega$$

RKHS H_k は以下で与えられる.

$$H_k = \left\{ f \in L^2(\mathbf{R}^m, dx) \,\middle|\, \int \frac{|\hat{f}(\omega)|^2}{\rho(\omega)} d\omega < \infty \right\}$$

$$\langle f, g \rangle = \int \frac{\hat{f}(\omega)\overline{\hat{g}(\omega)}}{\rho(\omega)} d\omega$$

ただし，$\hat{f}(\omega)$ は f のFourier変換

$$\hat{f}(\omega) = \frac{1}{(2\pi)^m} \int \exp\left(-\sqrt{-1}\omega^T z\right) f(\omega) d\omega$$

証明の概要:

– $\langle f, g \rangle$ はH_k の内積を定める(自明)

– $k(\ , x)$ はH_k の再生核となる.

$$k(x,y) = \int \exp\left(\sqrt{-1}\omega^T (x-y)\right)\rho(\omega)d\omega$$
$$= \int \exp\left(\sqrt{-1}\omega^T x\right)\boxed{\exp\left(-\sqrt{-1}\omega^T y\right)\rho(\omega)}d\omega$$

⇨ $k(\ , y)$ (y fixed)のFourier変換は

$$\exp\left(-\sqrt{-1}\omega^T y\right)\rho(\omega)$$

⇨ $$\langle f, k(\cdot, y)\rangle = \int \frac{\hat{f}(\omega)\overline{\exp(-\sqrt{-1}y^T\omega)\rho(\omega)}}{\rho(\omega)}d\omega$$
$$= \int \hat{f}(\omega)\exp(\sqrt{-1}y^T\omega)d\omega = f(y)$$

RKHSのFourier変換による表示

- Gauss Kernel

$$k_G(x,y)=\exp\left(-\frac{1}{2\sigma^2}(x-y)^2\right) \qquad \rho(\omega)=\frac{1}{2\pi}\exp\left(-\frac{\sigma^2}{2}\omega^2\right)$$

$$H_k=\left\{f\in L^2(\mathbf{R}^m,dx)\;\middle|\;\int|\hat{f}(\omega)|^2\exp\left(\tfrac{\sigma^2}{2}\omega^2\right)d\omega<\infty\right\}$$

$$\langle f,g\rangle=2\pi\int\hat{f}(\omega)\overline{\hat{g}(\omega)}\exp\left(\tfrac{\sigma^2}{2}\omega^2\right)d\omega$$

- Laplacian Kernel

$$k_L(x,y)=\exp(-\beta|x-y|) \qquad \rho(\omega)=\frac{1}{2\pi(\omega^2+\beta^2)}$$

$$H_k=\left\{f\in L^2(\mathbf{R}^m,dx)\;\middle|\;\int|\hat{f}(\omega)|^2(\omega^2+\beta^2)d\omega<\infty\right\}$$

$$\langle f,g\rangle=2\pi\int\hat{f}(\omega)\overline{\hat{g}(\omega)}(\omega^2+\beta^2)d\omega$$

セクション2のまとめ

- 正定値カーネル
  - 特徴写像の内積を与えるカーネル　$\langle \Phi(x), \Phi(y) \rangle = k(x, y)$
  - 例：　多項式カーネル，ガウスカーネル，ラプラスカーネル
- 再生核ヒルベルト空間
  - 再生核を持つ関数空間
  - 正定値カーネルと再生核ヒルベルト空間は1対1に対応
- 特徴空間・特徴写像

$$\Phi : \Omega \to H_k, \quad \Phi(x) = k(\cdot, x)$$

  を考えれば十分
- 基本的性質
  - 正定値カーネルを保つ演算：　和，積，極限，正規化
  - Bochnerの定理：　平行移動不変な正定値カーネルのFourier表示

参考文献

Aronszajn, N. (1950). Theory of Reproducing Kernels. Transactions of the American Mathematical Society 68 (3): 337–404.

Berlinet, A. and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in Probability and Statistics*, Kluwer Academic Publishers, 2004.

福水 「カーネル法入門　ー 正定値カーネルによるデータ解析」 朝倉書店（2010）